Accurate Molecular Electrostatic Potentials Based on Modified PRDDO/M Wave Functions: II. Electrostatic Potentials Inside the Molecular van der Waals Envelope

DENNIS S. MARYNICK

Valerian Software, 3058 Creekview Drive, Grapevine, Texas 76051

Received 24 October 1996; accepted 21 April 1997

ABSTRACT: In part I of this series, the *PESP* (parameterized electrostatic potential) method was described and applied to the calculation of electrostatic-potential-derived charges for a wide variety of organic and inorganic systems. Based on PRDDO/M wave functions and parameterized against *ab initio* MP2/6-31G** calculations, *PESP* is an order of magnitude faster than *ab initio* STO-3G calculations, while achieving a level of accuracy that rivals that of far more sophisticated *ab initio* methods. In this study, the application of the *PESP* method to the high potential regions of molecules containing H, C, N, O, F, P, S, Cl, and Br is described. For a collection of 48 molecules and 55 distinct lone pair minima, *PESP* yields the location and depth of lone pair minima to an average accuracy (relative to MP2/6-31G**) of 0.03 Å and 2.5 kcal/mol, respectively. Similarly, the location and well depths of minima in the π regions of organic molecules are calculated to an accuracy of 0.08 Å and 1.5 kcal/mol. *PESP* electrostatic potential maps are, in some cases, virtually indistinguishable from those obtained at the MP2/6-31G** level. © 1997 John Wiley & Sons, Inc. *J Comput Chem* **18**: 1682–1693, 1997

Keywords: electrostatic; potential; PESP; PRDDO

Correspondence to: D. S. Marynick. E-mail: valerian@ix.netcom.com

CCC 0192-8651 / 97 / 131682-12

Introduction

It is now well-recognized that electrostatic potentials (ESPs) play an important role in the analysis and understanding of long range noncovalent interactions.[2–21] Molecular electrostatic potential (MEP) minima often correlate with proton affinities and give information on the relative stability of tautomeric forms of organic species.[2] They have also proved useful in the analysis of steric/electronic aspects of nucleophilic attack,[5] and play a crucial role in self-consistent reaction field solvation theories.[21] Many other potential applications of MEPs have been discussed in the literature, including applications to crystal packing effects, protein-ligand binding, catalysis, surface electrostatics, and especially ESP-derived atomic charges.[10–15] A number of excellent reviews are available.[3, 5d, 6, 9]

In part I of this series,[1] a new approach to the calculation of electrostatic potential derived atomic charges was presented. Based on modified PRDDO/M[22–27] wave functions, the *PESP* (parameterized electrostatic potential) approach employs a minimum basis set of Slater orbitals but is parameterized against *ab initio* MP2/6-31G** calculations. For a collection of 820 atoms in 145 molecules containing H, C, N, O, F, P, S, Cl, and Br, including hypervalent species for P, S, and Cl, *PESP* achieved an overall average absolute deviation (unscaled) of $0.037e^-$, with a correlation coefficient of 0.990. Intermediate in computational complexity between a fully *ab initio* and a fully semiempirical approach, the *PESP* method is an order of magnitude faster than the simplest *ab initio* approach (STO-3G) while achieving a degree of accuracy approaching that of the reference calculations which it is parameterized against. As implemented in part I, *PESP* achieves high accuracy through a direct parameterization of the wave function and analytic (but rapid) calculation of the resultant electrostatic potentials.

In this study, a modification of the *PESP* method will be described, which makes it applicable to the high potential regions of molecules (inside the molecular van der Waals envelope). In this region, the ESP cannot be calculated from ESP-derived atomic charges, but must be calculated directly from the wave function. This is due to the fact that ESP-derived charges represent the atom-centered monopole approximation to the molecular charge distribution, and yield the correct ESP only in the limit of large distances. The modifications described here yield electrostatic potentials inside the molecular van der Waals envelope to a high degree of accuracy, while retaining the high accuracy found previously for electrostatic-potential-derived atomic charges (which are derived from sampling the electrostatic potential well outside the van der Waals envelope). The accuracy of the *PESP* method in high potential regions is documented by the calculation of the position and depth of local minima in the electrostatic potential associated with lone pairs and π bonding regions. For a collection of 48 molecules and 55 distinct lone pair minima, *PESP* yields the location and depth of lone pair minima to an average accuracy (relative to MP2/6-31G**) of 0.03 Å and 2.5 kcal/mol, respectively. Similarly, the location and well depths of minima in the π regions of organic molecules are calculated to an accuracy of 0.08 Å and 1.5 kcal/mol. This performance is competitive with *ab initio* HF/6-31G*-level calculations, while requiring about two orders of magnitude less computing time. *PESP* electrostatic potential maps are, in some cases, virtually indistinguishable for those obtained at the MP2/6-31G** level.

The remainder of this article is structured as follows. First, the PRDDO/M method and the *PESP* modifications to the PRDDO/M wave function are briefly reviewed. Then the modifications to the *PESP* method which are necessary to achieve accuracy in the high potential regions are presented, along with an outline of the parameterization strategy. Detailed statistics for the location and magnitude of lone pair and π region minima are then presented. Then, a detailed comparison of electrostatic potential maps for four molecules (adenine, formamide, uracil, and 1,1-chlorofluoroethylene) is presented. Finally, an outline of future extensions of the *PESP* method will be presented.

PRDDO/M APPROACH AND *PESP* MODIFICATIONS

Partial retention of diatomic differential overlap (PRDDO[22–24]) is an approximate *ab initio* technique designed to reproduce an *ab initio* minimum basis set calculation in a fraction of the computational time. PRDDO/M[25–27] is a modern version of the method, for which calculations in the

500–1500 orbital range are routine on workstations, and calculations approaching 4000 orbitals have been performed[26] on supercomputers. Both PRDDO and PRDDO/M are parameterized against *ab initio* calculations (as opposed to experiment) and they retain a significant number of two-electron integrals. With a few exceptions,[27] PRDDO employs a minimum basis set of Slater orbitals. It performs calculations in a basis of orthogonal atomic orbitals (OAOs), which are obtained via a Löwdin symmetric orthogonalization[28] of the input basis set. In the OAO basis, integrals of the form $(\chi_i \chi_j | \chi_k \chi_l)$ are vanishingly small when $i \neq j \neq k \neq l$ and may be neglected without significant loss of accuracy.[22] Thus, only an N^3 set of integrals must be computed. In principle, a PRDDO calculation is nothing more than a Hartree–Fock calculation in the OAO basis. However, because an exact OAO transformation would require the complete N^4 set of integrals in the Slater basis, approximate transformations must be employed. In some cases, these approximately transformed integrals are then adjusted via parameterizations against a *reference parameter set* (RPS) of *ab initio* calculations. PRDDO/M introduces three important modifications of the original PRDDO method: (1) a more sophisticated parameterization is performed; (2) valence-electron-only calculations using modified frozen core potentials (FCPs) are allowed[26]; and (3) calculations can be performed in a basis which is almost, but not quite, orthogonal (the not quite orthogonal atomic orbital, or NQOAO basis).[25] The NQOAO option makes very large calculations possible with essentially no loss in accuracy, especially with regard to the charge distribution. The PRDDO/M method employs four parameters for each atom (five for the atoms Al–Cl with a valence *d* orbital). These parameters modify integrals in the OAO basis. Further details of the method may be found in the original studies.[25–27]

The essence of the *PESP* modifications to the PRDDO/M method is extremely simple. A *PESP* calculation is simply a PRDDO/M/FCP calculation with a new set of parameters chosen to minimize the error function:

$$\varepsilon = \sum_i |q_i^{ab\ initio} - q_i^{prddo}| \qquad (1)$$

where q_i is the electrostatic-potential-derived atomic charge for atom i. Most of the modified parameters in the *PESP* approach are just orbital exponents. Only two of the PRDDO/M/FCP parameters (p_1 and p_3 in Ref. 1) are modified, and then only for some atoms. To obtain electrostatic-potential-derived atomic charges, the *PESP* wave function is analytically backtransformed to the input Slater orbital basis, the electrostatic potential is calculated accurately[1] for a set of points outside the van der Waals envelope of the molecule,[29] and standard fitting procedures[12, 13, 29] are employed to obtain the atomic charges.

Unlike some other attempts at parameterizing electrostatic potentials,[31, 32] the *PESP* calculations described in Ref. 1 do not parameterize the evaluation of the electrostatic potential, but rather the wave function itself. Even though the *PESP* wave function is derived from a minimum basis set calculation, the asymptotic behavior of the resultant charge distribution is essentially correct.

PESP Calculations in Regions of High Electrostatic Potential

GENERAL CONSIDERATIONS

Although the *PESP* wave function produces accurate ESPs outside the van der Waals envelope of a molecule, it is significantly less accurate in regions of high electrostatic potential. This is clearly a result of the use of a minimum basis set. For instance, analytic evaluation of the ESP at the *ab initio* STO-6G level on ammonia predicts the location and depth of the lone pair minimum at 1.00 Å and −110.5 kcal/mol, respectively. The corresponding values obtained at the MP2/6-31G** level are 1.22 Å and −83.8 kcal/mol. *PESP*, as defined in Ref. 1, yields 1.00 Å and −116.4 kcal/mol. Similar trends exist for nitrogen in more complicated molecules.[31] It seems clear that the accurate calculation of ESPs inside the van der Waals envelope of molecules will require modifications of the *PESP* approach. Such modifications must satisfy the following criteria: (1) they must result in calculated ESPs inside the van der Waals envelope that are in much better agreement with *ab initio* MP2/6-31G** results; and (2) they must not adversely affect the accuracy of the electrostatic potentials outside the van der Waals envelope, which are used to obtain ESP-derived atomic charges. Such an approach is described here.

The basic approach, which is similar to one successfully employed by others,[31,32] is to modify the calculation of the electrostatic potential itself. In the spirit of the approximate *ab initio* nature of PRDDO/M, the electrostatic integrals are first calculated accurately, and then they are modified by a function that contains adjustable parameters derived by comparison to *ab initio* ESPs. Although all three-center ESP terms are calculated,[1] only two-center terms are modified. Specifically, for the evaluation of the ESP at the point C, only the integrals $(ns_A np_A|1/r_C)$ and $(np_A np_A|1/r_C)$ are modified. Here, n represents the valence principle quantum number. After considerable experimentation, the functional form of the parameterized two-center ESP integrals for *sp* integrals was taken as:

$$\left(ns_A np_A|\frac{1}{r_c}\right)_{PARA} = \left(ns_A np_A|\frac{1}{r_c}\right)_{ab\ initio} \times \left(1 + \frac{C_{sp}E_{sp}^{-r}}{\left(ns_A np_A|\frac{1}{r_c}\right)_{ab\ initio}}\right) \quad (2)$$

An analogous equation was assumed for *pp* integrals. This modification is applied to all nonzero components of the two-center *sp* and *pp* integrals in the local diatomic coordinate system (before rotation to the general coordinate system). For hydrogen, the $(1s_A 1s_A|1/r_C)$ integrals are modified in a similar fashion.

COMPUTATIONAL DETAILS

To determine parameters C and E for each atom, a large number of points was generated for a set of molecules using the Merz–Kollman algorithm[12,13,29,30] with artificially short van der Waals radii. Typically, points were generated in the region 0.7 R_{vdw} to 1.4 R_{vdw} which sampled the ESP in a relatively uniform and unbiased manner. Points with *ab initio* ESPs greater than 100 kcal/mol were eliminated. In addition, points around lone pair minima and points in the π regions of organics were usually added to the data sets to give somewhat more statistical weight to the regions around local minima of the electrostatic potential. Parameters for hydrogen, carbon, nitrogen, and oxygen were optimized simultaneously, and then the remaining atoms (fluorine, chlorine, phosphorus, sulfur, and bromine) were parameterized individually in that order. The molecules used to determine the parameters are listed in Table I, along with a summary of the statistical results. *Ab initio* ESPs were calculated at the MP2/6-31G** level with the program GAUSSIAN-94[33] on an NEC SX-3 computer. *PESP* ESPs were calculated analytically by procedures described earlier[1] subject to the modifications introduced in eq. (2). The geometries for all molecules were taken from Ref. 1 or, if not contained in Ref. 1, were generally optimized at the HF/6-31G* (6d) level with the program GAMESS.[34] Geometry optimizations were performed on an IBM RS-6000 Model 250 computer. A simplex procedure was used for optimization of the parameters. During the optimization procedure, a penalty function was applied if the correction to the analytically calculated two-center ESP *sp* and *pp* integrals exceeded 0.002 a.u. (0.001 a.u. for phosphorus) at a distance of 1.4 times the van der Waals radius of the nearest atom. This insured that the correction factors [eq. (2)] died off quickly and did not affect the calculation of electrostatic-potential-derived atomic charges.

Results and Discussion

A total of 50 molecules and 8214 ESP evaluations in the range of 98.3 to −83.6 kcal/mol were used to define the parameters. For this data set, *PESP* achieves an overall average deviation of 2.5 kcal/mol relative to MP2/6-31G** electrostatic potentials (Table I). As demonstrated in Table II, the method is uniformly accurate to 2–3 kcal/mol in the regions of negative ESP, but is slightly less accurate in regions of very high ESP. This is clearly due to the intentional overemphasis of the negative ESP regions in the statistical fitting procedure. The final values for the parameters [eq. (2)] are listed in Table III.

LONE PAIR MINIMA

In Table IV, the location and well depths of lone pair minima for 36 minima in 29 organic molecules are shown, along with the corresponding HF/6-31G* and MP2/6-31G** values. Sixteen of these

TABLE I.
Molecules Included in Parameterization and Overall Statistical Results.

Data set	Molecules	No. of points	Average deviation[a]	Range[a]
CHNO	C_2H_2, C_2H_4, C_2H_6, CH_2NH, CH_2O, CO, CH_4, CO_2, CH_3NH_2, ethylene oxide, H_2O, formic acid, HCN, HNO, N_2, NH_2OH, NH_3, pyridine	2018	2.9	97.1 to −83.6
F	1,1-Difluoroethylene, HF, OF_2, NF_3, FCH_3, CF_2O, fluorobenzene	1244	2.5	98.3 to −31.9
Cl	1,1-Dichloroethylene, chlorobenzene, $ClCH_3$, CCl_4, ClF_3, $HClO_4$, NCl_3, OCl_2, phosgene	1139	2.6	68.8 to −27.2
P	PH_3, $P(CH_3)_2(OCH_3)$, $P(OH)_3$, $(C_2H_3)PH2$, phosphole	1478	2.3	65.8 to −48.7
S	H_2S, $(CH_3)_2S$, CH_3SNH_2, CH_3SOH, thiophene, $(C_2H_3)SH$	1461	3.0	74.1 to −48.6
Br	1,1-Dibromoethylene, BrCCH, bromopyridine, CH_3Br, bromobenzene	874	1.5	64.5 to −59.1
All atoms	All molecules	8214	2.5	98.3 to −83.6

[a]kcal / mol.

TABLE II.
Summary of Errors in Calculated ESPs for Molecules in Table I.

ESP range[a]	Number of points	Average absolute error[a]
0 / 10	1163	2.6
10 / 20	1587	2.7
20 / 30	598	3.4
30 / 40	320	3.8
40 / 50	231	4.1
50 / 60	107	4.3
60 / 70	80	4.9
70 / 80	16	6.2
80 / 90	9	4.8
90 / 100	13	6.2
0 / −10	1259	2.1
−10 / −20	1177	2.1
−20 / −30	951	1.9
−30 / −40	314	1.9
−40 / −50	172	1.9
−50 / −60	93	2.4
−60 / −70	92	2.8
−70 / −80	162	2.5
−80 / −90	12	2.2

[a]kcal / mol.

minima are associated with molecules not in the data set used to define the parameters. (For the numbering system employed for adenine and uracil, see Fig. 1). Table V presents similar data for inorganic systems. Table VI presents the statistical results for lone pair minima. For organics, *PESP* errors in the distances of the lone pair minimum to

TABLE III.
Two-Center Parameters.

Atom	C_{ps}	E_{ps}	C_{pp}	E_{pp}
Hydrogen[a]	0.4558	9.5151	—	—
Carbon	0.8160	4.5609	−18.6249	43.1730
Nitrogen	1.1242	5.8625	4.5573	21.2875
Oxygen	0.7780	5.8402	2.1135	18.1647
Fluorine	0.4697	4.9406	1.0215	5.3415
Phosphorus	−0.0062	1.4677	1.7961	5.1984
Sulfur	−0.2775	2.9283	1.9000	6.2816
Chlorine	0.2722	6.3776	1.2417	6.7257
Bromine	−5.9352	3.7207	2.6904	6.7207

[a]For hydrogen, the tabulated numbers refer to C_{SS} and E_{SS}.

TABLE IV.
Lone Pair Minima for Organic Systems.[a]

	PESP		MP2[b]		HF[c]	
	Dist.	Value	Dist.	Value	Dist.	Value
Pyrizole[d]	1.24	−65.6	1.26	−65.1	1.20	−65.1
Adenine[d]						
(N_6)	1.22	−60.2	1.26	−56.9	1.24	−62.3
(N_{12})	1.23	−64.1	1.26	−59.2	1.23	−66.8
(N_{15})	1.21	−64.7	1.26	−59.7	1.23	−67.8
Formaldehyde	1.24	−40.6	1.27	−40.0	1.23	−49.2
Pyridine	1.22	−69.6	1.26	−65.3	1.23	−70.8
Ethylene oxide	1.22	−41.1	1.25	−41.1	1.22	−53.7
Nitromethane[d]	1.18	−77.3	1.22	−79.3	1.21	−84.3
Formic acid						
(O)	1.29	−26.8	1.30	−26.9	1.27	−29.2
(═O)	1.23	−43.4	1.27	−41.7	1.23	−49.8
Methanol[d]	1.21	−51.6	1.21	−55.3	1.19	−60.6
Furan[d]	1.31	−28.3	1.31	−28.8	1.26	−36.5
Formamide[d]						
(O)	1.21	−55.7	1.20	−51.7	1.26	−61.6
(═O)	1.20	−59.3	1.20	−55.5	1.20	−65.1
Uracil						
(O_1)	1.21	−49.3	1.25	−48.3	1.21	−57.3
(O_1)	1.23	−46.5	1.26	−45.5	1.21	−54.6
(O_6)	1.24	−43.8	1.27	−42.2	1.23	−50.3
(O_6)	1.23	−46.3	1.26	−43.9	1.22	−52.4
Phosgene (O)	1.35	−25.2	1.40	−18.8	1.34	−26.4
CF_2O (O)	1.38	−19.6	1.37	−23.4	1.32	−29.3
Methyl fluoride	1.22	−28.7	1.24	−29.7	1.21	−35.2
Fluorobenzene	1.30	−18.3	1.31	−17.7	1.27	−22.8
Methyl chloride	1.81	−16.4	1.81	−18.1	1.78	−20.4
Chlorobenzene	1.89	−14.0	1.88	−15.1	1.83	−18.0
Ethyl chloride[d]	1.78	−19.0	1.80	−19.4	1.77	−22.3
$P(CH_3)_2(OCH_3)$	1.70	−37.8	1.76	−36.7	1.72	−40.6
$P(OH)_3$	1.70	−20.8	1.77	−18.7	1.74	−19.0
Phosphole	1.84	−26.0	1.89	−26.3	1.82	−32.5
Ethyl phosphine[d]	1.79	−33.3	1.81	−35.9	1.78	−38.9
Dimethyl sulfide	1.69	−32.8	1.72	−35.1	1.69	−38.4
Vinyl sulfide	1.79	−20.9	1.83	−22.5	1.79	−25.9
Ethyl bisulfide[d]	1.74	−29.6	1.74	−32.6	1.72	−35.4
Thiophine	1.97	−15.4	2.04	−12.8	1.91	−16.4
Bromobenzene	2.05	−14.8	2.03	−14.8	1.98	−17.1
Bromomethane	1.99	−15.2	1.96	−17.1	1.92	−19.0
Parabromopyridine	2.14	−7.4	2.10	−7.7	2.06	−8.9

[a]Distances (Å) and value (kcal / mol).
[b]6-31G** basis.
[c]6-31G* basis.
[d]Not used for parameterization.

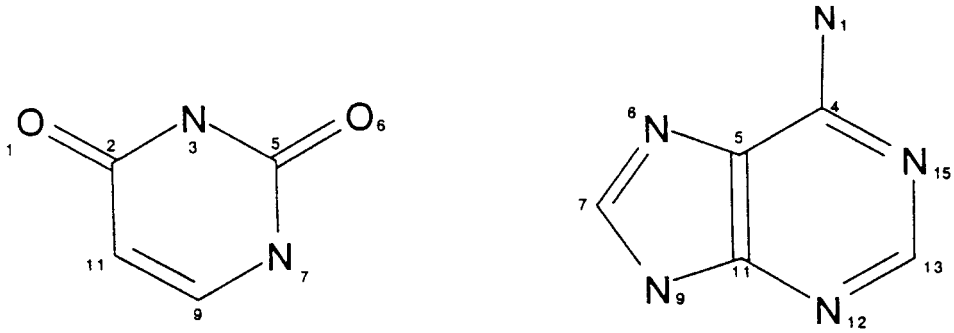

FIGURE 1. Numbering system for uracil (left) and adenine (right).

the atom associated with the lone pair are, on average, smaller than those found at the HF/6-31G* level (*PESP*: 0.028 Å; HF: 0.040 Å), although the latter errors are more systematic. Similarly, *PESP* yields well depths that are significantly more accurate that those calculated at the Hartree–Fock level (*PESP*: 2.0 kcal/mol; HF: 5.6 kcal/mol). Beyond the quantitative accuracy, the qualitative trends also seem reasonable. For instance, *PESP* yields the correct order of the well depths for all three minima in adenine and all four minima in uracil.

For inorganics (Table V) *PESP* is somewhat less accurate, although it is still quite competitive with HF/6-31G* calculations for the molecules examined here. Triple-bonded molecules, such as CO and N_2, are particularly troublesome. Indeed, elimination of these two molecules reduces the average errors in distances and well depths to 0.031 Å and 2.1 kcal/mol, respectively. A summary of the statistical errors in the location and depth of the lone pair minima for all molecules is given in Table VI.

TABLE V.
Lone Pair Minima for Inorganic Systems.[a]

	PESP		MP2[b]		HF[c]	
	Dist.	Value	Dist.	Value	Dist.	Value
CO (C)	1.44	−38.2	1.48	−22.3	1.49	−17.5
N_2	1.40	−22.3	1.50	−13.1	1.53	−10.8
Ammonia	1.18	−81.8	1.22	−83.8	1.21	−88.8
HCN	1.31	−48.6	1.35	−46.4	1.33	−47.4
HNO (N)	1.32	−34.0	1.36	−32.3	1.37	−30.5
NH_2OH (N)	1.21	−73.9	1.24	−75.8	1.23	−76.1
NCl_3 (N)	1.33	−29.7	1.37	−27.3	1.34	−29.9
NF_3 (N)	1.49	−7.4	1.47	−9.3	1.48	−5.1
Water	1.19	−60.6	1.20	−59.4	1.19	−63.0
CO (O)	1.54	−7.0	1.54	−6.5	1.45	−14.6
HNO (O)	1.21	−40.5	1.28	−34.5	1.26	−38.5
CO_2	1.57	−12.0	1.48	−14.8	1.44	−17.7
NH_2OH (O)	1.25	−56.9	1.23	−56.3	1.22	−59.4
OCl_2 (O)	1.32	−19.1	1.36	−17.8	1.34	−21.3
F_2	1.56	−1.5	1.44	−4.0	1.47	−3.5
ClF_3 (F)	1.26	−20.4	1.29	−21.6	1.25	−26.0
HF	1.22	−30.4	1.22	−32.3	1.21	−34.6
PH_3	1.87	−24.8	1.85	−29.0	1.83	−30.4
H_2S	1.75	−22.9	1.78	−27.3	1.77	−28.8

[a]Distances (Å) and value (kcal/mol).
[b]6-31G** basis.
[c]6-31G* basis.

π MINIMA

Another molecular region of interest is the π region of organic systems. The calculation of the location of minima in the π regions is especially difficult, because the electrostatic potential varies extremely slowly. This is illustrated in Figure 2, where the ESPs for ethylene, benzene, pyridine,

TABLE VI.
Summary of Average Absolute Errors in Distances (ε_D) and Potentials (ε_V) for Lone Pair Minima.[a]

	ε_D *PESP*[b]	ε_D HF[b,c]	ε_V *PESP*[d]	ε_V HF[c,d]
Organics	0.028	0.040	2.0	5.6
Inorganics	0.043	0.023	3.3	2.9
Overall	0.032	0.034	2.5	4.6

[a]Relative to MP2/6-31G** values.
[b]Angstroms.
[c]6-31G* basis.
[d]Kilocalories per mole.

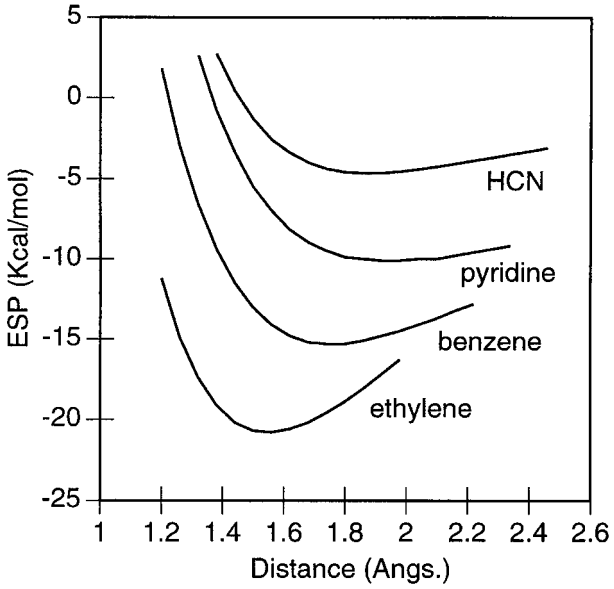

FIGURE 2. Variation of electrostatic potential in the π regions of simple multiple bonded species (see text).

and HCN at the MP2/6-31G** level along a line bisecting the multiple bond and perpendicular to the molecular plane are presented. For benzene, the electrostatic potential varies by less than 0.15 kcal/mol in the region ± 0.1 Å from the minimum. This variation is even smaller for pyridine and many other heterocyclic aromatics. Table VII illustrates the quantitative comparison of *PESP*, MP2/6-31G**, and HF/6-31G* data in the π regions. *PESP* achieves an average error in distances and well depths of 0.08 Å and 1.5 kcal/mol, respectively. The corresponding errors at the HF level are 0.05 Å and 2.2 kcal/mol. Again, the HF errors are more systematic, especially for distances.

ELECTROSTATIC POTENTIAL MAPS

Figures 3–7 illustrate *PESP* and MP2 electrostatic potential maps for a variety of molecules not included in the parameterization. For uracil (Fig. 3), adenine (Fig. 4), foramide (Fig. 5), and 1,1-chlorofluoroethylene (Fig. 6), two maps are shown, one in the molecular plane with contour levels of ± 10 kcal/mol and one 1.5–2.0 Å above the plane, with contour levels of ± 3 kcal/mol. For uracil, adenine, and formamide, the *PESP* contour maps are virtually indistinguishable from their MP2 counterparts. However, 1,1-chlorofluoroethylene shows some clear differences in the plot 1.5 Å above the

TABLE VII.
Minima in π Regions of Organic Systems.[a]

	PESP		MP2[b]		HF[c]	
	Dist.	Value	Dist.	Value	Dist.	Value
Benzene	1.66	−17.7	1.76	−15.3	1.73	−19.0
Acetylene	1.46	−25.4	1.54	−21.0	1.51	−25.4
Ethylene	1.50	−19.5	1.54	−20.8	1.52	−25.6
HCN	1.75	−4.4	1.87	−4.7	1.76	−7.4
Pyridine						
C — N	1.83	−10.8	1.93	−10.1	1.86	−13.0
C_o — C_m	1.86	−7.2	1.95	−6.5	1.88	−9.5
C_m — C_p	1.82	−6.1	1.86	−6.1	1.83	−9.3
Chlorobenzene						
C_1 — C_1	2.10	−10.5	2.24	−8.2	2.21	−9.7
C_1 — C_2	1.83	−11.3	1.99	−8.2	1.93	−10.6
C_2 — C_3	1.79	−9.5	1.85	−8.2	1.81	−11.2
C_3 — C_4	1.75	−10.4	1.83	−8.9	1.79	−11.3
Parafluoropyridine[d]						
C_2 — C_1	2.12	−1.9	2.05	−2.5	1.98	−3.9
C_1 — N	1.98	−5.1	1.96	−6.4	1.94	−7.6
Furan[d]						
O — C_1	1.88	−10.5	1.97	−9.3	1.93	−11.4
C_1 — C_2	1.58	−17.5	1.68	−15.7	1.64	−19.0
C_2 — C_3	1.69	−15.2	1.71	−16.3	1.72	−17.9
Phenol[d]						
C — O	1.78	−12.4	1.86	−10.3	1.90	−11.4
C_1 — C_2	1.77	−10.8	1.89	−10.8	1.85	−13.2
C_2 — C_3	1.66	−13.8	1.74	−13.9	1.73	−16.2
C_3 — C_4	1.66	−17.1	1.74	−15.1	1.72	−17.7
Pyrizole[d]						
N — N_H	1.73	−11.1	1.80	−12.7	1.70	−17.5
N_H — C_1	1.80	−7.0	1.88	−7.9	1.85	−10.0
C_1 — C_2	1.56	−17.9	1.68	−15.7	1.65	−18.3
C_2 — C_3	1.64	−18.4	1.71	−17.3	1.72	−18.9
C_3 — N	1.64	−17.8	1.74	−18.1	1.69	−21.6
Adenine[d]						
N_1 — C_4	1.88	−6.9	1.86	−9.8	1.97	−9.7
C_4 — N_{15}	2.10	−7.8	1.99	−11.0	1.99	−11.8
C_{13} — N_{15}	1.98	−9.8	1.87	−13.3	1.88	−13.1
N_{12} — C_{13}	1.90	−9.6	1.85	−13.2	1.87	−12.8
C_{11} — N_{12}	1.97	−7.2	1.95	−9.7	1.95	−10.5
C_5 — C_{11}	2.10	−5.6	2.18	−6.6	2.11	−8.0
N_9 — C_{11}	2.34	−2.4	2.24	−3.4	2.34	−4.0
N_6 — C_7	1.77	−6.8	1.85	−7.4	1.75	−10.2
C_5 — N_6	1.82	−9.6	1.90	−9.5	1.79	−13.4
C_4 — C_5	2.20	−6.4	2.29	−7.3	2.23	−8.6

[a] Minimum of the electrostatic potential along a line perpendicular to the molecular plane intersecting the midpoint of the indicated bond (Å and kcal/mol). The average absolute errors (Å and kcal/mol) relative to MP2/6-31G** are: for distances, 0.08 (*PESP*) and 0.05 (HF); for potentials, 1.5 (*PESP*) and 2.2 (HF).
[b] 6-31G** basis.
[c] 6-31G* basis.
[d] Molecules not used for parameterization.

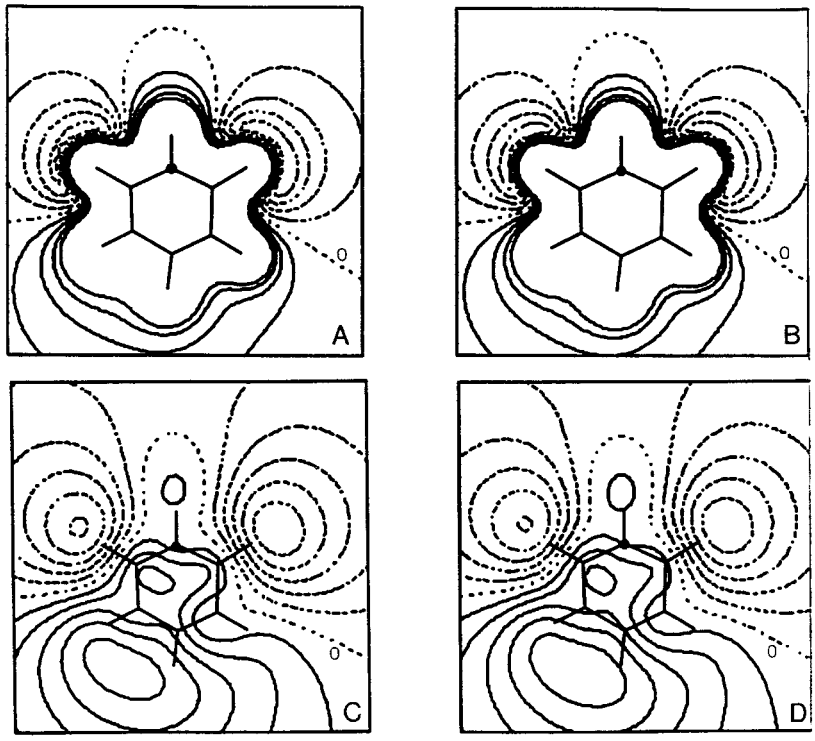

FIGURE 3. Electrostatic potential maps for uracil. (A) MP2 / 6-31G** and (B) *PESP* in the molecular plane. Contour level ±10 kcal / mol. (C) MP2 / 6-31G** and (D) *PESP* 2 Å above the molecular plane. Contour level ±3 kcal / mol.

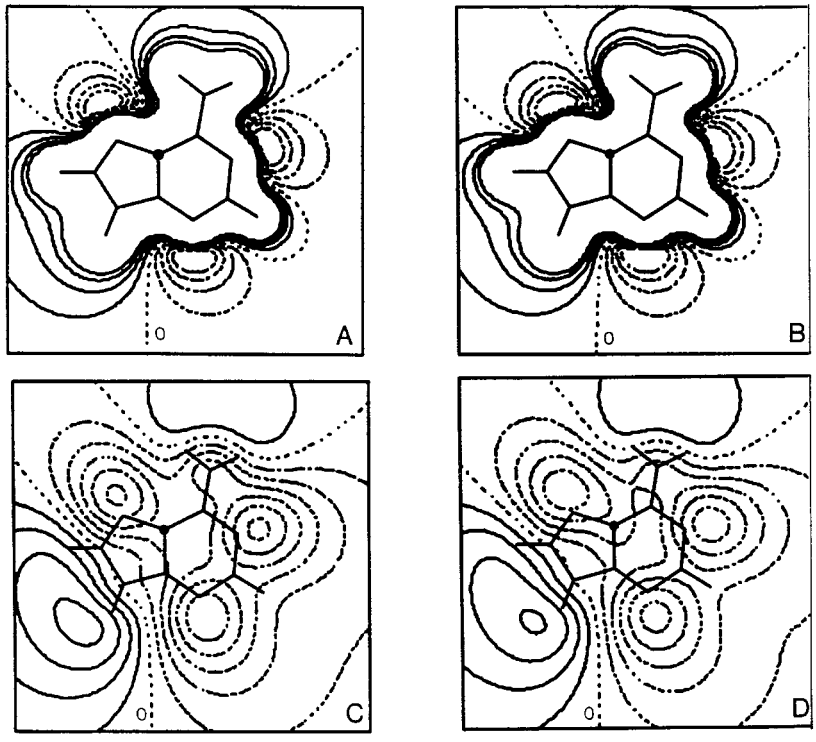

FIGURE 4. Electrostatic potential maps for adenine. (A) MP2 / 6-31G** and (B) *PESP* in the molecular plane. Contour level ±10 kcal / mol. (C) MP2 / 6-31G** and (D) *PESP* 2 Å above the molecular plane. Contour level ±3 kcal / mol.

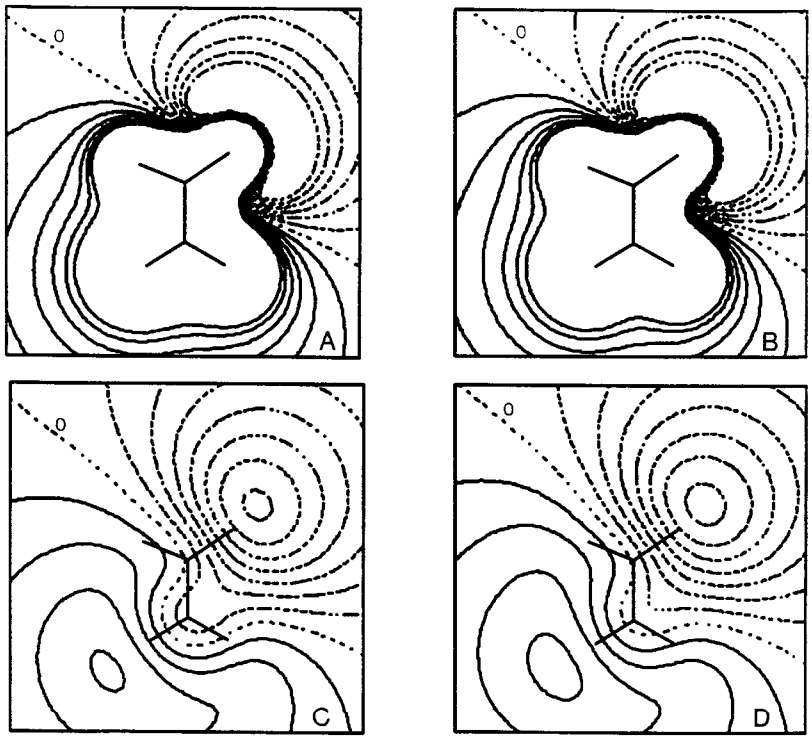

FIGURE 5. Electrostatic potential maps for formamide. (A) MP2 / 6-31G** and (B) *PESP* in the molecular plane. Contour level ±10 kcal / mol. (C) MP2 / 6-31G** and (D) *PESP* 2 Å above the molecular plane. Contour level ±3 kcal / mol.

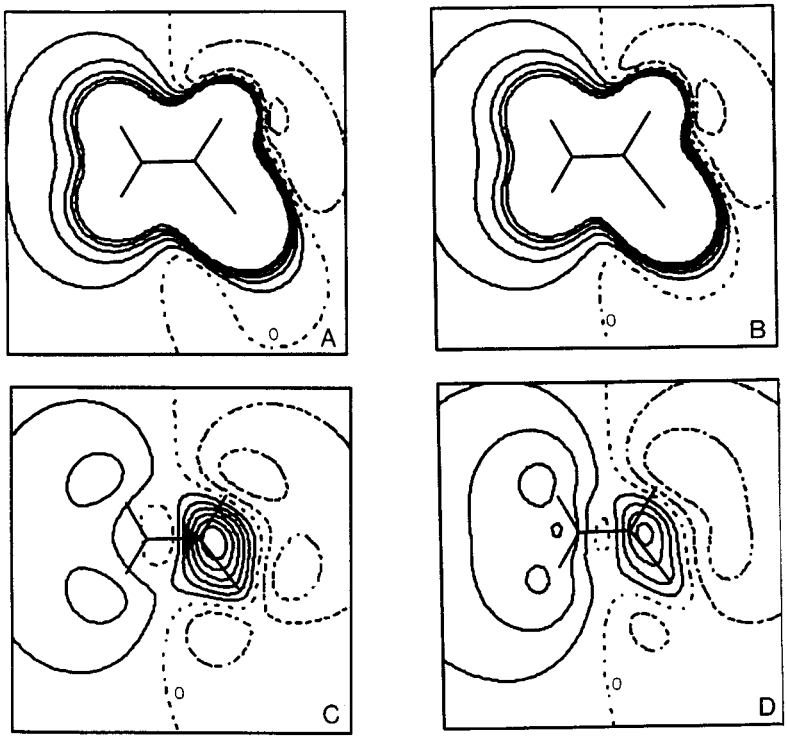

FIGURE 6. Electrostatic potential maps for 1,1-chlorofluoroethylene. (A) MP2 / 6-31G** and (B) *PESP* in the molecular plane. Contour level ±10 kcal / mol. (C) MP2 / 6-31G** and (D) *PESP* 1.5 Å above the molecular plane. Contour level ±3 kcal / mol.

molecular plane, mainly in the area over C_1. Here, *PESP* is underestimating the (positive) electrostatic potential by about 6 kcal/mol. Figure 7 illustrates the electrostatic potential for adenine in a plane perpendicular to the molecule. Again, the agreement between *PESP* and MP2 is excellent.

EFFECT OF NEW PARAMETERIZATION ON ELECTROSTATIC-POTENTIAL-DERIVED ATOMIC CHARGES

Although the parameterization described here is designed to have minimal effect on the evaluation of the ESP in regions well outside the van der Waals envelope of the molecule, the new parameters will affect the ESP-derived charges. It is therefore of interest to examine the effect of the two-center parameterizations introduced here on the ESP-derived charges. Table VIII summarizes the results. For the 820 unique atoms in 145 molecules described in part I of this series, the new parameterization has essentially no statistical effect on the calculated charges. The largest changes in average absolute charge error (relative to MP2/6-31G**) occur for carbon ($0.005e^-$ decrease) and phosphorous ($0.004e^-$ increase). The overall average absolute error remains $0.037e^-$ for the entire 820-atom data set.

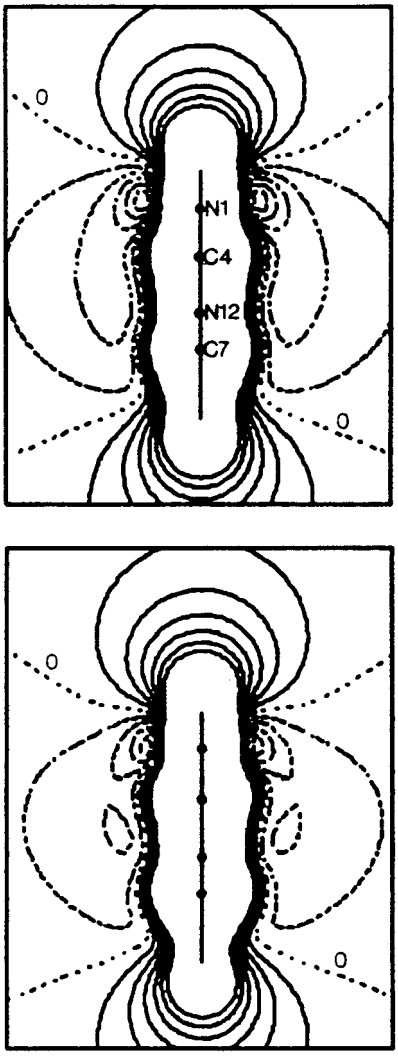

FIGURE 7. Electrostatic potential maps for adenine in a plane perpendicular to the molecule. Top: MP2/6-31G**; bottom: *PESP*. Contour level ±3 kcal/mol.

Conclusions

Simple modifications in the evaluation of the ESP of *PESP* wave functions produce accurate electrostatic potentials in regions of high potential, while retaining the high accuracy previously found for the calculation of electrostatic-potential-derived atomic charges. The *PESP* approach is about an order of magnitude faster than the simplest of *ab initio* calculations[1] (about two orders of magnitude faster that HF/6-31G*), but is clearly competitive in accuracy to the latter method. The next article in this series will deal with the extension of the method to other atoms. Preliminary parameters are already available for Na^+ and K^+, and parameterizations are planned for Mg^{++}, Ca^{++}, and Zn^{++}. Ultimately, it is easy to foresee further improvements in the method. Simultaneous opti-

TABLE VIII.
ESP-Derived Charges from the Original *PESP* Method and This Work.

Atom	Number of unique atoms	Average abs. charge	Average unsigned error[a]	Average unsigned error[b]
H	299	.205	.019	.019
C	218	.325	.062	.057
N	62	.641	.048	.048
O	109	.452	.033	.032
F	35	.148	.026	.027
P	21	.422	.064	.068
S	22	.380	.069	.069
Cl	33	.194	.042	.042
Br	21	.098	.030	.030
Overall	820	.307	.037	.037

[a] Ref. 1 (original *PESP* parameterization).
[b] This work.

mization of the high and low potential parameters may well improve agreement in both regions. Additionally, parameters can easily be developed for specific chemical environments (such as the backbone atoms of polypeptides and sugars). Further articles in this series will deal with these issues.

Acknowledgments

The author thanks the Swiss Center for Scientific Computing and Dr. Djordje Maric for a generous grant of CPU time required for the calculation of the *ab initio* ESPs at the MP2 level.

References

1. D. S. Marynick, *J. Comput. Chem.* **18**, 955 (1997).
2. E. Scrocco and J. Tomasi, *Adv. Quantum Chem.*, **11**, 115 (1978).
3. P. Politzer and J. S. Murray, In *Reviews of Computational Chemistry*, *Vol. 2*, K. B. Lipkowitz and D. B. Boyd, Eds., VCH, New York, ch. 7.
4. P. Politzer and D. G. Truhlar, Eds., *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, Plenum Press, New York, 1981.
5. (a) H. B. Bürgi, *Angew. Chem.*, **87**, 461 (1975); (b) J. Seres, G. Nàray-Szabó, K. Simon, K. Daróczi-Csuka, I. Szilágyi, and L. Párkányi, *Tetrahedron*, **37**, 1565 (1981); (c) A. T. Pudzianowski, J. C. Barrish, and S. H. Spergel, *Tetrahed. Lett.*, **37**, 293 (1992); (d) G. Náray-Szabó, *Chem. Design Automat. News*, **8**, 43 (1992).
6. J. G. Ángyán and G. Náray-Szabó, In *Theoretical Models of Chemical Bonding, Part 4, Theoretical Treatment of Large Molecules and Their Interactions*, Z. B. Maksic, Ed., Springer, Berlin, 1991, p. 1.
7. J. E. Douglas and P. A. Kollman, *J. Am. Chem. Soc.*, **102**, 4295 (1980).
8. G. Pepe, D. Siri, and J.-P. Reboul, *J. Mol. Struct. Theochem*, **14**, 289 (1981).
9. H. Weinstein, R. Osman, and J. P. Green, In *Computer Assisted Drug Design*, E. C. Olson and R. E. Christoffersen, Eds., *ACS Symposium Series 112*, American Chemical Society, Washington, DC, 1979.
10. F. A. Momany, *J. Phys. Chem.*, **85**, 592 (1978).
11. S. R. Cox and D. E. Williams, *J. Comput. Chem.*, **2**, 304 (1981).
12. U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **5**, 129 (1984).
13. L. M. Chirlian and M. M. Francl, *J. Comput. Chem.*, **8**, 894 (1987).
14. C. M. Breneman and K. B. Wiberg, *J. Comput. Chem.*, **11**, 361 (1990).
15. C. I. Bayly, P. Cieplak, W. D. Cornell, and P. A. Kollman, *J. Phys. Chem.*, **97**, 10269 (1993).
16. M. M. Francl, *J. Phys. Chem.*, **89**, 428 (1985).
17. P. Politzer and M. Levy, *J. Chem. Phys.*, **87**, 5044 (1987).
18. M. Orozco and F. F. Laque, *J. Comput. Chem.*, **14**, 587 (1993).
19. P. Politzer, P. R. Laurence, and K. Jayasurlya, *Environ. Health. Perspect.*, **61**, 191 (1985).
20. J. S. Murry and P. Politzer, *Chem. Phys. Lett.*, **152**, 364 (1988).
21. S. Miertus, E. Scrocca, and J. Tomasi, *Chem. Phys.*, **55**, 117 (1981).
22. T. A. Halgren and W. N. Lipscomb, *J. Chem. Phys.*, **58**, 1569 (1973).
23. T. A. Halgren and W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA*, **69**, 652 (1972).
24. D. S. Marynick and W. N. Lipscomb, *Proc. Natl. Acad. Sci. USA*, **79**, 1341 (1982).
25. A. Derecskei-Kovacs and D. S. Marynick, *Int. J. Quantum Chem.*, **58**, 193 (1996).
26. A. Derecskei-Kovacs, D. E. Woon, and D. S. Marynick, *Int. J. Quantum Chem.*, **61**, 67 (1997).
27. A. Derecskei-Kovacs and D. S. Marynick, *Int. J. Quantum Chem.* **63**, 1091 (1997).
28. P.-O. Löwdin, *J. Chem. Phys.*, **18**, 365 (1950).
29. M. L. Connolly, *J. Appl. Cryst.*, **16**, 540 (1983).
30. B. H. Besler, K. M. Merz, Jr., and P. A. Kollman, *J. Comput. Chem.*, **11**, 431 (1990).
31. G. P. Ford and B. Wang, *J. Comput. Chem.*, **14**, 1101 (1993).
32. B. Wang and G. P. Ford, *J. Comput. Chem.*, **15**, 200 (1994).
33. M. J. Frisch, G. W. Trucks, H. B. Schlegel, P. M. W. Gill, B. G. Johnson, M. A. Robb, J. R. Cheeseman, T. Keith, G. A. Petersson, J. A. Montgomery, K. Raghavachari, M. A. Al-Laham, V. G. Zakrzewski, J. V. Ortiz, J. B. Foresman, C. Y. Peng, P. Y. Ayala, W. Chen, M. W. Wong, J. L. Andres, E. S. Replogle, R. Gomperts, R. L. Martin, D. J. Fox, J. S. Binkley, D. J. Defrees, J. Baker, J. P. Stewart, M. Head-Gordon, C. Gonzalez, and J. A. Pope, *GAUSSIAN-94*, Gaussian, Inc., Pittsburgh, PA, 1995.
34. Original program: M. Dupuis, D. Spangler, and J. J. Wendoloski, Program QG01, National Resource for Computations in Chemistry, Software Catalog, University of California, Berkeley, CA (1980). This version: M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.*, **14**, 1347 (1993).